
Accroche-toi au niveau, j'enlève l'échelle

Eléments d'analyse des aspects multi-niveaux dans la simulation à base d'agents

J. Gil-Quijano* — G. Hutzler* — T. Louail*, **

* Laboratoire IBISC (Université Evry-Val d'Essonne)
523 Place des Terrasses de l'Agora, 91000 Evry
quijano@ibisc.fr; hutzler@ibisc.fr; thomas.louail@ibisc.fr

** Laboratoire Géographie-Cités, UMR 8504 (CNRS, Universités Paris 1-Paris 7)
13 rue du Four, 75006 Paris

RÉSUMÉ. Lorsque l'on modélise des systèmes complexes, il est souvent nécessaire de considérer des entités à plusieurs niveaux d'organisation et niveaux d'échelles. La prise en compte de ces niveaux, de leurs influences réciproques, et des dynamiques d'organisation aux interfaces entre eux, est un problème ardu pour lequel les solutions proposées sont souvent liées à une discipline ou à un cas d'étude particulier. Pour essayer d'aller vers une méthodologie plus générale de conception de modèles multi-niveaux, nous proposons une grille d'analyse des approches utilisées, élaborée notamment à partir de l'étude de trois exemples en biologie et en géographie. Nous montrons ensuite, dans le cadre d'un exemple unificateur, comment différentes approches peuvent être combinées.

ABSTRACT. Modelling complex systems often implies to consider entities at several levels of organisation and levels of scales. Taking into account these levels, their mutual interactions, and the organizational dynamics at the interface between levels, is a difficult problem, for which the proposed solutions are often linked to a specific disciplinary field or a particular case study. In order to develop a broader methodology for designing multilevel models, we propose an analytical framework of existing approaches, drawn in particular from the study of three examples in biology and geography. We finally show, through an unifying example, how different approaches can be combined.

MOTS-CLÉS : simulation à base d'agents, dynamiques multi-niveaux, simulation multi-niveaux, biologie cellulaire, dynamiques urbaines, colonies de fourmis

KEYWORDS: agent-based simulation, multilevel dynamics, multilevel simulation, cellular biology, urban dynamics, ant colonies

1. Introduction

Les systèmes complexes naturels et sociaux sont souvent caractérisés par de nombreuses entités hétérogènes, de nature et de dimensions diverses, à différents niveaux d'organisation, interagissant de manière complexe les uns avec les autres, à des niveaux d'échelles de temps et d'espace eux aussi très divers. Dans ce cadre, l'aspect « multi-niveaux » d'une modélisation correspond le plus souvent au choix d'un niveau particulier de description, et à l'observation de structures émergentes de plus haut niveau. Le choix d'un niveau conditionne la sélection des entités du système réel qui sont réifiées sous forme d'agents dans le modèle. Le système simulé peut alors être analysé à au moins deux niveaux d'organisation : celui des agents d'une part, via une analyse de l'évolution de leur état, et celui du système d'autre part, via des mesures permettant de caractériser les structures produites au cours de la simulation. Le système peut aussi être mesuré à de multiples niveaux intermédiaires qui peuvent faire sens pour le modélisateur, correspondant par exemple à des *groupes* d'agents partageant des caractéristiques communes, même si en pratique ce raffinement est plus rarement mis en place.

Il nous semble dans ce cas difficile de parler de modélisation véritablement multi-niveaux dans la mesure où seul le niveau « du bas » est présent dans le modèle et spécifié par le modélisateur. Le ou les niveaux supérieurs sont dans ce cas des niveaux qui peuvent être observés par l'expérimentateur qui « regarde » et mesure les simulations, mais ils ne sont pas réifiés en tant que tels dans le modèle. Le choix du niveau de modélisation du système est particulièrement important car il doit satisfaire un certain nombre de contraintes qui s'imposent au modélisateur, sous peine de ne pas être pertinent. Ce dernier doit en particulier s'interroger sur la question à laquelle son modèle est censé répondre, sur les ontologies manipulées par les experts du domaine, sur les données disponibles aux différents niveaux de description, et enfin sur la puissance de calcul impliquée par le nombre d'entités et la période temporelle à simuler.

Cette approche émergentiste, mono-niveau dans la conception et bi-niveaux dans l'analyse des comportements produits, est de loin la plus courante actuellement en simulation orientée agents (voir par exemple les modèles présentés dans (Treuil *et al.*, 2008)). Elle n'est pourtant pas toujours applicable de manière satisfaisante pour des systèmes réels. En effet, un certain nombre de systèmes complexes ne peuvent se comprendre sans intégrer plusieurs niveaux d'organisation et des dynamiques qui ne peuvent se réduire à une vision purement émergentiste. Les systèmes sociaux en particulier intègrent de manière forte un rôle structurant des niveaux supérieurs sur les niveaux inférieurs. De ce fait, les théories du domaine correspondantes, permettant de décrire le fonctionnement du système, sont rarement mono-niveau, soit par manque de connaissance à un niveau donné, soit par le fait que les ontologies manipulées mixent des entités et des acteurs de niveaux d'organisation différents. Il est par exemple difficile d'imaginer reproduire l'évolution d'une ville en partant de la seule spécification du comportement de ses habitants. Par ailleurs, il serait trop coûteux d'un point de vue computationnel de simuler l'ensemble d'un système de grande taille au niveau le plus fin : simuler au niveau atomique, ou même au niveau moléculaire, l'ensemble

d'une cellule constituée de plus de 4.10^{13} molécules (cellule hépatique de rat) est ainsi parfaitement irréaliste. Il faut noter enfin que les données collectées sont souvent fragmentaires et hétérogènes, ce qui implique que l'on possède rarement les données permettant de caractériser entièrement l'état et la dynamique d'un système à un niveau particulier.

La question posée nécessite finalement la prise en compte d'entités à différents niveaux (moléculaire et cellulaire par exemple, ou ceux des individus, du quartier et de la ville) ce qui, comme le rappellent Servat et ses collègues dans (Servat *et al.*, 1998), fait partie de la gymnastique intellectuelle scientifique : faire cohabiter des points de vue et de description complémentaires d'un même système. De fait, notre perception de la réalité et des objets qui nous entourent se fait à plusieurs niveaux d'échelles. Ces auteurs pointent deux raisons supplémentaires à la prise en compte de niveaux d'organisation multiples. La première est d'ordre informatique : beaucoup de systèmes distribués font preuve d'un comportement global qui devient remarquable lorsque le nombre de composantes interagissant devient critiquelement grand. Substituer des « super-agents » à des groupes d'agents et leur associer un modèle comportemental plus global mais équivalent permet d'une part de gagner en performance et d'autre part de transposer les ressources ainsi économisées pour modéliser plus finement d'autres parties du système. La seconde raison avancée est d'ordre plus thématique. La prise de conscience de sa similarité avec d'autres agents et de l'existence d'une structure « groupe » les rassemblant peut entraîner chez l'agent une modification intentionnelle de son comportement, soit pour renforcer son appartenance au groupe, soit pour s'en distinguer. Plus généralement, le fait de pouvoir prendre en compte plusieurs niveaux d'organisation dans un même modèle doit permettre d'aborder des questions originales par rapport à un modèle classique à deux niveaux, celui de la spécification et celui des structures collectives émergentes.

Il semble donc important de pouvoir intégrer, au sein d'un même modèle, les descriptions d'entités de différents niveaux sur les échelles spatiale et temporelle, correspondant à différents niveaux d'organisation du système. Cependant et comme le souligne Lesne dans (Lesne, 2009), « il ne s'agit pas tant de prendre en considération tous les éléments et leurs relations que de considérer *simultanément* le niveau des éléments et ceux de leurs propriétés émergentes et la façon dont ils sont couplés ». Pour le formuler autrement, c'est donc plus dans l'*articulation* de différents niveaux d'organisation que dans leur seule juxtaposition que se situe l'apport d'un modèle multi-niveaux. Articuler les niveaux, cela implique de les prendre en compte explicitement dans le modèle et de définir la manière dont ils sont couplés, ce qui se traduira dans un modèle à base d'agents par l'interaction entre agents/objets de niveaux différents. Cette question de l'articulation est d'actualité en simulation orientée agent. Treuil et ses collègues la présentent comme l'un des « chantiers » du domaine en conclusion de leur livre (Treuil *et al.*, 2008), et des projets actuels visent à construire des plates-formes de simulation devant faciliter la prise en compte simultanée de plusieurs niveaux (par exemple (Gignoux *et al.*, 2005) en écologie, la plate-forme Magéo en géographie).

Avant d'illustrer ces différents aspects par la présentation de trois exemples de modèles multi-niveaux dans la partie 3, nous commençons dans la partie 2 par un travail de définition indispensable pour préciser les notions manipulées. A partir des exemples présentés, pour lesquels nous explicitons les différents niveaux du modèle ainsi que les mécanismes assurant leur couplage, nous proposons une abstraction des organisations multi-agents sous-jacentes pour mettre en lumière leurs différences et discuter des conditions de leur déploiement. Ceci se traduit dans la partie 4 par une grille d'analyse de différents types d'approches multi-niveaux. Nous mettons ensuite en évidence les potentialités d'application de ces différentes approches à d'autres questions de modélisation ainsi que leur complémentarité en discutant leur utilisation dans un modèle de super-colonie de fourmis. Nous concluons en discutant des perspectives de recherche ouvertes par notre approche comparative.

2. Multi-échelles ou multi-niveaux ?

Echelle ou niveau, multi-échelles ou multi-niveaux ? Les deux notions d'échelle et de niveau sont fréquemment utilisées pour modéliser les systèmes complexes et caractériser les abstractions manipulées, leur granularité et leur position dans une hiérarchie qui structure et organise le système. Les deux termes sont fréquemment utilisés de manière interchangeable, qui plus est de manière différente en fonction des disciplines, ce qui est potentiellement source de grande confusion. Il apparaît donc nécessaire de s'arrêter un instant sur ces deux notions, mais aussi sur celles de *hiérarchie* et de *structure*. Nous nous référerons largement, pour cela, au travail effectué par Gibson et ses collègues dans (Gibson *et al.*, 2000).

2.1. *Echelles et niveaux*

La notion d'*échelle* est fréquemment précisée par un qualificatif la rapportant à l'espace ou au temps : échelle spatiale, échelle temporelle. Le terme de *niveau* est quant à lui souvent associé à des qualificatifs tels que *micro*, *macro* ou encore *méso* situant le système ou le phénomène décrit en liaison avec sa taille, son temps d'évolution caractéristique ou un rapport d'inclusion avec d'autres systèmes ou sous-systèmes. En première approche, nous considérerons ainsi que la notion d'*échelle* se rapporte à une dimension d'analyse selon laquelle le phénomène d'intérêt pourra être mesuré. Cette dimension peut être spatiale ou temporelle, mais également quantitative. Les dimensions spatiale et temporelle se réfèrent respectivement à la taille des entités impliquées dans le phénomène (typiquement du nanomètre au millier de kilomètres en passant par toutes les tailles intermédiaires), et au temps caractéristique associé aux comportements de ces entités et à leurs interactions (de la nanoseconde au siècle). La dimension quantitative se réfère quant à elle au nombre d'entités impliquées dans le phénomène (typiquement de 2 à quelques milliards).

Par rapport à cette notion d'échelle, la notion de *niveau* sera utilisée pour situer le phénomène étudié, et/ou les entités qui le composent, le long de la dimension d'ana-

lyse utilisée. Un niveau correspondra ainsi généralement à l'ensemble des entités dont les dimensions ou le temps d'évolution caractéristique sont d'un ordre de grandeur identique ou comparable. Si l'on s'en tient à cette utilisation des notions d'échelle et de niveau, il ne faudrait donc pas, en toute rigueur, parler de modèles multi-échelles. Il s'agit en effet, la plupart du temps, de signifier que les entités du modèle se situent à des niveaux différents le long de l'échelle spatiale. Nous préférons pour notre part parler de *modèles multi-niveaux* et nous ne parlerons donc plus de *modèles multi-échelles* dans la suite de l'article.

Il faut noter immédiatement que la notion de niveau est à la fois *relative* et liée à un choix de modélisation. Par *relative*, nous entendons qu'une entité considérée pour la description d'un phénomène ne peut être décrite de manière absolue comme étant de niveau micro, méso ou macro. Du point de vue de la biologie moléculaire par exemple, les molécules sont associées au niveau micro tandis que la cellule est associée au niveau macro. Du point de vue de la physiologie, les cellules sont au contraire associées au niveau micro, le niveau macro correspondant au tissu. Par ailleurs, le fait d'associer une entité à un niveau n'est pas uniquement lié à ses caractéristiques le long des dimensions d'échelles mais constitue le résultat d'un choix d'analyse et de modélisation. Le choix d'associer différentes entités au sein d'un niveau fait en effet souvent intervenir le fait que les entités appartiennent à une structure ou à une organisation pertinente du point de vue de la description du système. Il est d'ailleurs fréquent de parler de « niveau d'organisation » ou de « niveau d'abstraction ».

On voit ainsi apparaître deux nouvelles notions qui sont celles de *hiérarchie* et de *structure*. La notion de hiérarchie, car les systèmes sont souvent décrits sous la forme d'un emboîtement hiérarchique de niveaux (molécules, cellules, tissus, organes, individus, groupe, société). La notion de structure, car c'est souvent la reconnaissance de la structuration d'un ensemble d'entités à un niveau donné qui permet la définition d'entités à un niveau supérieur. (Gibson *et al.*, 2000) présentent la notion de *hiérarchie* comme « un système de regroupement d'objets ou de processus liés conceptuellement ou causalement le long d'une échelle analytique », et distinguent trois types de hiérarchies : exclusives, inclusives ou constitutives. Une *hiérarchie exclusive* est une hiérarchie dans laquelle il n'y a pas de lien d'inclusion entre les entités d'un niveau et les entités du niveau supérieur. C'est le cas par exemple des hiérarchies de commandement militaire ou encore des chaînes trophiques dans lesquelles des individus d'un certain niveau se nourrissent des individus de niveau inférieur. *A contrario*, dans les hiérarchies de types inclusive ou constitutive, les entités d'un niveau donné sont incluses dans une entité de niveau supérieur. Dans le cas des *hiérarchies inclusives*, il s'agit d'une relation catégorielle à la manière des hiérarchies de classes dans un langage de programmation objet ou bien des hiérarchies de catégories taxonomiques (domaine, règne, embranchement, classe, ordre, famille, genre, espèce). Dans le cas des *hiérarchies constitutives*, les entités d'un niveau sont regroupées en nouvelles entités au niveau immédiatement supérieur, ces dernières étant caractérisées par de nouvelles organisations, fonctions et propriétés émergentes (molécules, cellules, tissus, organes, individus). En plus des niveaux d'échelles des entités considérées, c'est la *structure*

du groupe qu'elles forment, c'est-à-dire la manière dont elles sont organisées au sein du groupe, qui permettra de les positionner à un niveau ou à un autre.

2.2. *Le point de vue disciplinaire*

A présent que les principales notions ont été précisées, il est intéressant de regarder comment celles-ci sont manipulées et utilisées dans différents domaines disciplinaires. Les objets manipulés par les uns et les autres sont d'évidence très différents. Les niveaux d'échelles sont différents mais ce qui distingue sans doute encore davantage les champs disciplinaires, c'est la manière d'articuler ces niveaux, en rapport avec la manière de hiérarchiser les entités d'un système.

2.2.1. *Niveaux atomique et moléculaire*

Les phénomènes qui se présentent aux niveaux atomique et moléculaire sont étudiés en utilisant des approches diverses par différentes disciplines comme la physique, la chimie et la biologie moléculaire. Une constante unificatrice de ces différentes approches est de considérer de manière implicite ou explicite des hiérarchies de type constitutive : un solide, un liquide, un gaz ou une cellule sont constitués de molécules, elles-mêmes constituées d'atomes, eux-mêmes constitués d'électrons, neutrons et protons, ces deux derniers étant constitués de quarks. L'approche généralement adoptée dans ce cadre consiste à modéliser les entités d'un niveau et à tenter d'expliquer les propriétés observées au niveau immédiatement supérieur par auto-organisation, émergence, ou plus simplement par accumulation des effets produits par chacune des entités. On peut ainsi en physique expliquer la notion de pression d'un gaz, au niveau macro, par le nombre de chocs par unité de surface, induits par le déplacement des particules, modélisé au niveau micro. De même en chimie où la cinétique d'une réaction peut être expliquée par le comportement de l'ensemble des entités qui y participent au niveau inférieur. Dans ces approches de type bottom-up ou ascendant, les structures lorsqu'il y en a, ne sont pas pré-existantes mais sont le résultat induit des interactions entre les entités du niveau inférieur.

2.2.2. *Du niveau moléculaire au niveau de l'individu*

On se place ici dans le cadre de la biologie, et plus précisément de la physiologie, c'est-à-dire l'étude de l'organisation mécanique des êtres vivants à partir des relations entre leurs composants (organes, tissus et cellules). On peut identifier, de même qu'aux niveaux inférieurs, des relations hiérarchiques de types constitutives : un individu est constitué d'organes, formés de tissus, eux-mêmes constitués de cellules. Il serait absurde cependant, de prétendre décrire le fonctionnement d'un individu en le découpant, par une approche purement réductionniste, en organes dont on étudierait le fonctionnement indépendamment les uns des autres. Que ce soit au niveau moléculaire ou au niveau des organes, le fonctionnement des différents éléments identifiés à un niveau donné est très fortement couplé, c'est-à-dire qu'on ne peut étudier un élément en l'isolant des autres mais qu'il faut au contraire, dans une approche

systemique, l'étudier dans ses interactions avec les autres éléments. Au niveau moléculaire, les gènes et les protéines ne peuvent pas être étudiés du simple point de vue de la cinétique chimique des réactions dans lesquelles ils sont impliqués mais ils doivent être appréhendés du point de vue des réseaux de régulation génique ou des réseaux métaboliques auxquels ils participent. Au niveau physiologique, ce sont les différents systèmes de l'organisme (nerveux, reproducteur, circulatoire, respiratoire, moteur, digestif, etc.) qui fonctionnent en étroite interaction les uns avec les autres. Par ailleurs, les différents niveaux sont loin d'être étanches, ce qui impose de s'intéresser à leurs influences mutuelles. Pour reprendre l'exemple des gènes, on ne peut comprendre la manière dont ils sont différenciellement exprimés au cours du temps et dans les différentes parties de l'organisme que si l'on prend en compte tous les phénomènes épigénétiques qui contrôlent cette expression, sous l'influence d'une multitude de paramètres environnementaux.

2.2.3. *Niveaux écologique et social*

Dans les cas de l'écologie et des sciences sociales, on considère souvent des entités en interaction au sein de hiérarchies constitutives (individu - groupe - société) mais également au sein de hiérarchies exclusives (chaîne trophique, individu - norme - institution). Pour la compréhension de ce type de systèmes, il est souvent nécessaire de manipuler des abstractions définissant des niveaux d'analyse qui dépendent à la fois du problème modélisé mais également de la discipline du modélisateur : par exemple, une société humaine peut être analysée sous l'angle économique, social ou politique, et les niveaux d'analyse que l'on peut alors considérer sont les individus, les ménages, les profils économiques, les groupes sociaux, les flux migratoires, etc. Dans ce type de système, il est rare de s'intéresser à la formation de structures de niveau supérieur, on considère souvent que ces structures existent et l'on définit la dynamique des interactions entre ces structures et les individus (on ne s'intéresse pas à la formation d'une société, mais aux relations entre celle-ci et les individus qui la constituent). On s'intéresse alors à la formation et à l'évolution de motifs spatio-temporels (phénomènes de ségrégation, flux, équilibres proie-prédateur, etc.) ou de motifs sociaux (réseaux sociaux, petits mondes, etc.) que l'on suppose être le résultat des interactions entre individus et structures de niveau supérieur.

3. De la cellule biologique à la « cellule urbaine » : trois exemples de simulation multi-niveaux

Maintenant que le contexte a été précisé, nous allons nous attarder sur trois projets de simulation à base d'agents dans lesquels les problématiques de modélisation multi-niveaux ont été abordées explicitement. Le premier exemple traite de la croissance d'une tumeur cancéreuse, abordée aux niveaux moléculaire et cellulaire. Le second traite des choix de logement des ménages dans la ville de Bogotá, traitée individuellement au niveau des ménages et des logements mais également au niveau de *groupes* de logements et de ménages. Enfin le troisième présente un modèle de morphogénèse urbaine, approchée aux niveaux inter-urbain et intra-urbain.

Nous ne présentons pas ici les modèles pour eux-mêmes et leurs résultats, mais uniquement pour la manière dont ils mettent en oeuvre la notion de modélisation multi-niveaux. Pour ce faire, et afin de faciliter la comparaison entre ces différents modèles, nous répondrons en particulier aux questions suivantes :

- 1) *Spécification des niveaux* : quels sont les différents niveaux identifiés, quels sont les types d'objets et/ou d'agents définis à chacun de ces niveaux, et comment ces types sont-ils définis (sont-ils spécifiés *a priori* ou bien sont-ils découverts de manière dynamique ?) ;
- 2) *Instanciation des objets* : comment l'instanciation des objets et/ou agents est-elle effectuée (celle-ci s'effectue-t-elle de manière statique en créant explicitement les agents du modèle, ou bien ceux-ci sont-ils créés de manière dynamique pendant la simulation en fonction du contexte ?) ;
- 3) *Couplage entre les niveaux* : de quelle manière les différents niveaux sont-ils couplés (comment interagissent les agents de niveaux différents ?).

3.1. *Modèle de migration de cellules cancéreuses*

Contexte et objectif

En matière de cancer, l'un des facteurs de mauvais pronostic est lié à l'apparition de métastases, c'est-à-dire de tumeurs secondaires. Nous nous intéressons pour notre part aux conditions micro-environnementales, autour d'une tumeur cancéreuse, pouvant conduire à l'échappement métastatique d'une cellule (Lepagnot *et al.*, 2009). En particulier, nous étudions l'impact de la molécule PAI-1 (une protéine), suspectée d'être à l'origine de la transformation morphologique des cellules permettant l'échappement (Maquerlot *et al.*, 2006). Cette molécule est produite par les cellules elles-mêmes et peut se fixer sur la matrice extra-cellulaire. Les cellules disposent de récepteurs pour fixer puis internaliser PAI-1. Lorsque cette dernière est accrochée à la matrice extra-cellulaire, présente tout autour de la tumeur, elle fournit alors aux cellules des points d'adhésion qui leur permettent de migrer à travers la matrice.

Modèle

Le modèle développé intègre, de manière individuelle, aussi bien les cellules cancéreuses que les molécules de PAI-1. La modélisation de la croissance de la tumeur est obtenue par un comportement de division cellulaire : une cellule peut se diviser pour donner deux cellules filles au bout d'un temps variable qui dépend de son accès aux nutriments. Ce comportement prolifératoire est couplé à un comportement de répulsion entre cellules : deux cellules trop proches se repoussent. Le modèle intègre par ailleurs la dynamique de production et d'internalisation de PAI-1 par les cellules. Pour ce faire, une cellule produit des molécules de PAI-1 avec un certain taux. Dans le même temps, elle peut internaliser le PAI-1 en suspension, ainsi que le PAI-1 matriciel, avec une probabilité directement proportionnelle à son nombre de récepteurs. L'accès aux nutriments définit trois états différenciés pour les cellules de la tumeur : une cel-

lule ayant un accès satisfaisant aux nutriments (couche externe de la tumeur) est active et possède à la fois les comportements de prolifération et production/internalisation de PAI-1 ; une cellule ayant un accès insuffisant aux nutriments nécrose et meurt (couche interne de la tumeur) ; les cellules de la couche intermédiaire reçoivent quant à elles suffisamment de nutriments pour produire et internaliser PAI-1 mais plus suffisamment pour proliférer ; ces dernières sont qualifiées de quiescentes.

Lorsque les tumeurs simulées atteignent quelques milliers de cellules, il faut alors gérer plusieurs centaines de milliers de molécules de PAI-1. On se trouve ainsi rapidement limités par la taille des tumeurs que l'on peut simuler. La solution proposée consiste à abstraire certains détails du modèle dans les zones où ils ne sont pas utiles, ce qui permet au choix de simuler des tumeurs plus grandes ou d'ajouter des détails dans les zones d'intérêt. La zone d'intérêt principal se situant à l'interface entre la couche externe de la tumeur et la matrice extra-cellulaire, nous avons donc proposé de remplacer les couches internes de la tumeur par un modèle agrégé, constitué à la fois de cellules et de molécules de PAI-1. Le coeur de la tumeur et son modèle orienté agents sont ainsi abstraits par un modèle global de calcul de flux entrant et sortant. Du point de vue spatial, ce modèle est délimité par l'ensemble des cellules nécrosées ou quiescentes, à l'exception d'une couche de cellules quiescentes qui constitue la bordure du modèle agrégé. Ces cellules étant immobiles ou très peu mobiles, cela permet de s'abstraire des mouvements de répulsion entre cellules, ceux-ci étant principalement dus au comportement de division cellulaire dans la couche de cellules actives. Le modèle agrégé intègre toutes les cellules ainsi que toutes les molécules de PAI-1 situées dans le périmètre ainsi défini. Il s'agit alors de calculer les interactions entre ce modèle agrégé et les cellules ou molécules situées à l'extérieur. Il est relativement aisé de déterminer les cellules et molécules qui devront être intégrées au modèle agrégé : dans le cas des cellules, il s'agit d'examiner lesquelles sont quiescentes et entièrement entourées d'autres cellules quiescentes ; concernant les molécules, il suffit d'examiner lesquelles, dans leur mouvement de diffusion aléatoire, vont percuter la frontière du modèle agrégé. Pour déterminer combien de molécules seront au contraire « relâchées » par ce modèle, il est nécessaire d'évaluer la proportion de cellules nécrosées par rapport aux cellules quiescentes, afin de déterminer le nombre de cellules produisant ou internalisant ces molécules. Il est ensuite aisé de mettre à jour la quantité de molécules de PAI-1 à l'intérieur du modèle agrégé, en considérant que l'activité d'internalisation et d'externalisation de ce modèle agrégé est directement proportionnelle à la quantité de cellules quiescentes qu'il contient. En assimilant les molécules diffusant à l'intérieur du modèle agrégé à un gaz, on peut alors calculer la « pression » à l'intérieur du modèle, et donc déterminer la quantité statistique de molécules de PAI-1 devant quitter le modèle agrégé. Celles-ci sont alors distribuées aléatoirement à l'extérieur du modèle agrégé, au voisinage immédiat de la frontière.

Agents et niveaux

Dans ce modèle coexistent des objets de tailles très différentes, des cellules d'une part, des molécules d'autre part. Les niveaux correspondent dans ce cas (voir figure 1) à ces deux types d'objets, lesquels sont définis *a priori*. L'instanciation s'effectue

en créant une cellule initiale, qui elle-même crée d'autres cellules par divisions successives, et produit et externalise des molécules. Les niveaux cellulaire et moléculaire sont couplés du fait de l'activité des cellules : les cellules produisent et internalisent des molécules de PAI-1. L'ensemble produit par la simulation, à un niveau d'organisation supérieur, un amas de cellules que l'on observe et que l'on désigne sous le nom de tumeur. Il n'y a pas jusque là de passage de niveau mais seulement l'interaction, dans un environnement commun, entre objets de niveaux différents sur l'échelle spatiale. Le niveau de la tumeur est quant à lui seulement observé et n'a aucun rôle en tant que tel dans la simulation.

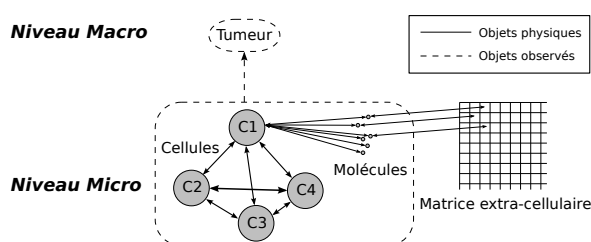


Figure 1. Modélisation d'une tumeur à partir d'entités cellules et molécules

A contrario, lorsque le modèle agrégé est introduit (voir figure 2), on identifie en cours de simulation une entité de niveau intermédiaire entre le niveau des cellules et le niveau de la tumeur qui est réifiée en tant que telle. Il y a donc l'introduction d'un niveau supplémentaire de modélisation, celui d'un tissu multicellulaire correspondant au coeur de la tumeur. Ce type d'objet est spécifié *a priori* : on sait de manière certaine que la simulation du modèle va conduire au développement d'une tumeur et l'on peut donc anticiper l'apparition de cet objet. Son instantiation en revanche dépendra du moment où un tel type d'objet pourra être observé dans la simulation. Concernant le couplage avec les cellules et les molécules, il s'effectue différemment selon que ces dernières sont internes ou externes au modèle agrégé. Le couplage avec les cellules et molécules internes est réalisé grâce à un système d'équations différentielles très simples, décrivant l'évolution des quantités de cellules respectivement nécrosées et quiescentes, l'évolution de la quantité de molécules à l'intérieur du modèle, ou encore la quantité de molécules externalisée. Le couplage avec les cellules et molécules externes est réalisé en décrivant les conditions dans lesquelles ces dernières doivent être intégrées au modèle agrégé, et les règles liées à l'externalisation de molécules.

3.2. SIMULBOGOTA : logements et ménages dans la ville de Bogotá

Contexte et objectif

L'objectif du modèle SIMULBOGOTA (Gil-Quijano *et al.*, 2007b) est de modéliser l'évolution de la distribution spatiale de la population de la ville de Bogotá sur plusieurs décennies. Cette évolution dépend principalement des migrations à l'intérieur

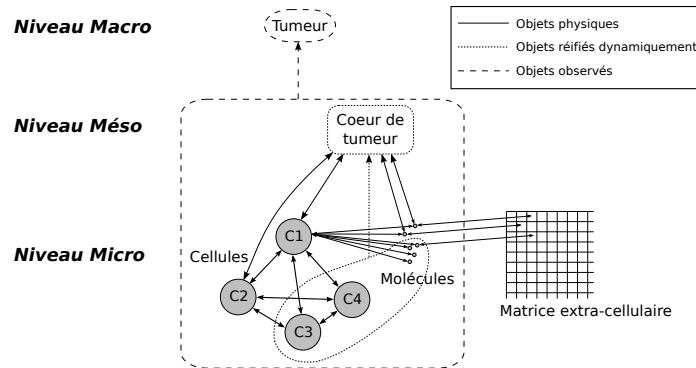


Figure 2. Introduction d'un modèle agrégé pour le cœur de la tumeur

et à l'extérieur de la ville ainsi que de l'évolution de la population de ménages et du parc de logements. L'évolution de la population de ménages est le résultat d'évènements individuels (mariage, divorce, émancipation, mort, etc.) qui produisent des changements socio-économiques dans les ménages existant, ainsi que la création et la disparition de ménages. L'évolution du parc de logements dépend principalement du renouvellement et du vieillissement du bâti. Dans le cas de Bogotá, ces processus sont le résultat de dynamiques complexes qui font intervenir la construction contrôlée par les politiques d'aménagement, mais également des mécanismes informels comme l'auto-construction et le lotissement illégal.

Modèle

Pour la réalisation de notre modèle, nous disposons uniquement de descriptions socio-économiques des ménages et des logements ainsi que de leur distribution spatiale pour les années 1973 et 1993. En l'absence de données suffisantes, une modélisation centrée sur une représentation explicite de mécanismes individuels de décision et d'évolution s'avère inadéquate voire impossible. Dans ces conditions, le parti pris adopté est de considérer des *groupes de ménages* et des *groupes de logements* comme entités principales de modélisation. Il s'agit de réduire la complexité des simulations par la prise en compte d'un niveau agrégé et d'y placer le processus de décision conduisant les ménages à déménager. Toutefois, il est nécessaire de considérer les populations de ménages et de logements afin de pouvoir modéliser l'évolution de ces groupes.

Dans ce modèle nous considérons de manière explicite deux niveaux de modélisation : le *niveau microscopique* qui contient les données qui décrivent la population de ménages et le parc de logements, et le *niveau mésoscopique* qui contient les groupes de ménages et les groupes de logements. Le modèle est composé essentiellement de trois mécanismes exécutés en séquence à chaque pas de simulation :

1) *formation de groupes* : cela permet de passer du niveau microscopique (ménages et logements) au niveau mésoscopique (groupes de ménages et groupes de logements). Il s'agit d'un mécanisme de classification automatique (Gil-Quijano *et al.*, 2007a) effectué sur les données qui représentent les ménages et les logements ;

2) *interaction entre groupes* : cela permet de relocaliser les ménages dans l'espace. Il s'agit d'un mécanisme basé sur des enchères qui permet l'échange de logements entre groupes de ménages. Ces interactions sont régies par des listes de préférences de logements construites dynamiquement¹ et par une matrice statique de coûts de déménagement entre secteurs urbains. Dans ce mécanisme, on fait l'hypothèse que les ménages cherchent à se rapprocher des ménages de leur groupe social et à occuper les mêmes groupes de logements. Il s'agit donc d'un mécanisme de ségrégation spatiale auto-renforcé ;

3) *évolution de la population* : il s'agit d'un mécanisme basé sur l'exécution de règles globales d'évolution de la population de ménages et de logements. Ces règles permettent la création ou la suppression d'entités au niveau microscopique (ménages et logements). Chaque règle associe un profil d'entité avec un nombre d'entités à traiter. Ces règles représentent des tendances globales d'évolution des effectifs, et non pas des événements individuels très difficiles à prendre en compte et très gourmands en données.

Pour pouvoir tracer l'évolution des groupes, un dernier mécanisme permet de mettre en relation les groupes trouvés au cours de deux pas de simulation successifs.

Agents et niveaux

Dans ce modèle a été introduit un niveau intermédiaire composé de structures « *artificielles* » : des groupes formés d'entités microscopiques semblables. Il s'agit de structures qui n'ont pas d'équivalent physique et qui résultent de l'analyse des données de ménages et de logements (niveau microscopique). On a donc la coexistence d'objets physiques au niveau micro avec des objets artificiels au niveau méso réifiés dynamiquement par classification des objets du niveau micro en groupes homogènes (voir figure 3). Par rapport aux modèles de phénomènes physiques, la question supplémentaire du niveau d'abstraction se pose : il s'agit de définir le nombre de groupes à considérer, sachant qu'un nombre réduit peut produire une sur-simplification du modèle, tandis qu'un nombre élevé peut rapprocher excessivement les groupes du niveau microscopique et entraîner la perte de leur capacité de synthèse. Même si le niveau des groupes a été *spécifié au préalable*, le nombre des groupes considérés est déterminé à l'exécution et les groupes eux-mêmes sont *instanciés dynamiquement* à partir de la classification des données de ménages et logements à chaque pas de simulation.

Du point de vue de l'évolution, des règles générales ont été construites à partir des données descriptives des ménages et des logements. Le moteur d'évolution est donc placé au niveau microscopique et ses effets sont propagés aux groupes par le

1. Les préférences dépendent de la distribution spatiale instantanée des ménages (*cf* (Gil-Quijano *et al.*, 2007b)).

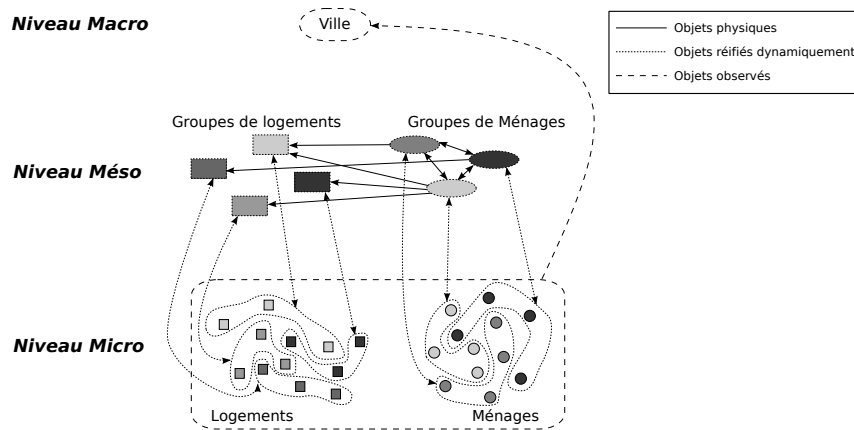


Figure 3. Diagramme des interactions entre niveaux dans la simulation de l'évolution de la distribution spatiale de la population dans SIMULBOGOTA ; Les flèches entre le niveau micro et le niveau méso correspondent aux processus de classification des ménages et logements en groupes de ménages et groupes de logements (dans le sens ascendant) et aux processus de relocalisation des ménages (dans le sens descendant)

mécanisme de passage entre le niveau microscopique et celui des groupes (couplage ascendant). En revanche, il n'existe pas d'interaction directe entre entités microscopiques, les entités *actives* du système étant donc les groupes. Le niveau microscopique fait office de *point d'ancrage avec la réalité* tandis que le niveau des groupes offre *une vision synthétique de la réalité*, ce qui facilite la formulation et l'évaluation d'hypothèses sur le comportement du système permettant donc de mieux comprendre le phénomène modélisé.

Le caractère « artificiel » des groupes conditionne la mise en place de mécanismes d'interaction *ad-hoc* entre groupes et des mécanismes d'interaction inter-niveaux. Contrairement à la modélisation d'entités « physiques », dans ce cas-là, et en l'absence de règles de comportements de groupe bien établies ou facilement déductibles, le modélisateur est en effet *libre* dans la définition de mécanismes d'interaction. Ceci d'autant plus que l'on connaît très peu sur le fonctionnement des groupes, ce qui implique que le mécanisme choisi pour représenter l'interaction entre les groupes est aussi artificiel que les groupes eux-mêmes. Le mécanisme proposé, basé sur les enchères de logement, n'a pas de fondement réel. Il s'agit d'un artefact qui permet l'appariement dynamique et multi-critères entre groupes de ménages et groupes de logements. Ce mécanisme, dont le choix ne relève que de l'intuition des modélisateurs, permet de relocaliser les ménages, et ainsi d'assurer le couplage descendant entre le niveau des groupes et le niveau microscopique.

3.3. SIMPOP3 : simuler la morphogénèse urbaine aux niveaux inter- et intra-urbain

Contexte et objectifs

Les modèles Simpop, conçus et développés au sein de l'équipe P.A.R.I.S du laboratoire Géographie-Cités, sont des modèles multi-agents de croissance urbaine sur le long-terme (plusieurs siècles). Le premier d'entre eux, le modèle SIMPOP1 (Bura *et al.*, 1996) fut l'une des toutes premières applications des SMA à la géographie. Dans cette section nous présentons deux modèles autonomes, SIMPOP2 et SIMPOP-NANO, puis leur couplage en SIMPOP3.

SIMPOP2² (Pumain *et al.*, 2009) est un projet démarré en 2002 qui donne une transcription calculatoire de la théorie évolutionnaire urbaine (Pumain, 1997), pour la mettre à l'épreuve de la simulation. La théorie analyse les systèmes de villes comme des systèmes complexes auto-organisés dont les villes sont les entités principales. Les systèmes de villes, que ce soit à l'échelle régionale, nationale ou continentale, présentent des propriétés remarquables au niveau macro, dont certaines sont particulièrement stables et universelles, comme la structure hiérarchique de la distribution des tailles de villes (connue sous le nom de loi rang-taille). En dissociant un modèle générique, qui encapsule les processus universels, des modèles appliqués, qui encapsulent les processus spécifiques à l'histoire de chacun de ces systèmes, la démarche comparative choisie permet de formaliser un ensemble de facteurs qui permettent de reproduire les différences observées dans les propriétés des systèmes de villes en Europe et aux Etats-Unis.

SIMPOP-NANO (Louail, 2009) fait suite à SIMPOP2 et a pour objectif de simuler l'évolution de l'organisation spatiale interne d'une ville, sur la même période pluri-séculaire que celle de SIMPOP2. Le but est de comparer les morphogénèses de villes en Europe et aux Etats-Unis, et d'isoler un ensemble minimal de facteurs permettant de reproduire les différences observées entre ces deux grands types de villes, en terme de motifs de répartition spatiale des densités, des prix et des activités à l'intérieur des villes des deux continents (Bertaud, 2004).

Avec SIMPOP3, l'objectif est de coupler les deux modèles précédents en un modèle, multi-agent et multi-niveaux, qui permette d'étudier des dynamiques urbaines simultanément aux niveaux inter et intra-urbain, et d'articuler les entités associées : le quartier, la ville, et le système de villes. L'idée sous-jacente est que ces niveaux exercent les uns sur les autres des influences réciproques. Etant donnée la multitude des facteurs causant ces influences inter-niveaux, il est difficile de les capturer et de formaliser leur action dans des règles inter-niveaux. Le modèle sert alors à tester des heuristiques sur la nature des inter-influences qu'entretiennent ces niveaux d'organisation du peuplement.

2. <http://www.simpop.parisgeo.cnrs.fr>

Le modèle SIMPOP2

Les *villes* sont les agents principaux du modèle. Elles interagissent en se vendant et en s'achetant des biens et des services associés aux *fonctions urbaines* qu'elles possèdent. Une *fonction urbaine* caractérise le rôle joué par une ville au sein du système, elle est associée à une période temporelle d'activité (fonction industrie lourde, industrie automobile, nouvelles technologies par exemple, mais aussi fonction capitale ou chef-lieu, etc.). Les interactions spatiales entre villes sont de plusieurs types, et dépendent de la *fonction urbaine* assurant l'interaction : proximité spatiale, délimitation spatiale administrative ou réseau privilégié à grande échelle. Dans le même temps, les villes sont en compétition pour l'acquisition d'innovations, représentées par de nouvelles fonctions urbaines qui apparaissent dynamiquement au cours de la simulation. À l'exécution, la dynamique interactionnelle engendre une spécialisation des villes, une différenciation de leurs trajectoires et donc la taille de leur population. La mesure détermine si ces structures émergentes possèdent les propriétés des systèmes de villes.

Le modèle SIMPOP NANO

Le modèle intègre deux types d'entités : des objets *quartiers*, qui représentent des portions de l'espace intra-urbain, et des agents *fonctions urbaines* qui représentent les types d'activités socio-économiques possédées par la ville. Ces *fonctions urbaines* sont les mêmes que celles qui sont présentes dans SIMPOP2, et sont dotées de comportements complémentaires pour trouver à s'exprimer à ce niveau intra-urbain. Les *fonctions* possèdent un effectif d'actifs à localiser dans la ville, ainsi qu'un budget permettant de payer pour leur localisation. Les quartiers sont reliés dans un réseau à structure dynamique et se différencient par leur accessibilité et par leur composition, qui les rendent plus ou moins attractifs pour les *fonctions*, et plus ou moins chers. Les *fonctions* se différencient par leurs budgets et leurs orientations synergiques. La dynamique des localisations de ces agents fonctions produit à l'exécution une ville *émergente*, dont il est possible de caractériser la structure spatiale interne au moyen de plusieurs indicateurs : gradients centre-périphérie de densité d'occupation du sol et de prix, coefficient de variation spatiale et coefficients de localisation, etc.

Le modèle SIMPOP3

SIMPOP3 inclut des agents à deux niveaux :

- Au niveau le plus « bas », niveau *micro*, sont spécifiés des *agents fonctions* qui sont les mêmes que ceux définis au niveau *méso* mais dotés de comportement supplémentaires leur permettant d'agir au niveau *micro*, c'est-à-dire de se localiser parmi les quartiers de la ville. Ce sont les agents de SIMPOP NANO.

- Au niveau intermédiaire, niveau *méso*, on spécifie des *agents villes*, capables d'interactions entre eux, ainsi que des *agents fonctions*, capables de se localiser parmi les villes, et qui offrent aux villes qui les possèdent de nouvelles capacités d'interaction. Ce sont les agents de SIMPOP2.

L'exécution du modèle consiste en l'exécution alternée d'une itération de SIM-

POP2 et d'une itération de SIMPOP_{NANO}.

- SIMPOP2 calcule à chaque pas de temps les échanges entre villes, la ventilation spatiale des innovations et la croissance résultante, et détermine en conséquence l'état suivant de chaque ville : une liste de fonctions, et l'effectif et le budget de chacune.

- Ces sorties sont les données d'entrée de SIMPOP_{NANO}, qui calcule la ventilation des effectifs des fonctions possédées par la ville au sein de ses quartiers. Les fonctions payent leur implantation dans les quartiers grâce à leur budget, calculé au niveau SIMPOP2, à la suite des échanges avec d'autres villes. En sortie d'itération, SIMPOP_{NANO} produit un indicateur de la qualité de l'implantation des fonctions à l'intérieur de la ville.

- Cet indicateur de « performance spatiale » est pris en compte par les agents-fonctions au niveau SIMPOP2 pour décider de la ville où elles vont s'implanter. La boucle inter-niveaux est bouclée. Nous modélisons à la fois des causalités descendantes et ascendantes.

Agents et niveaux

Dans SIMPOP2 le niveau spatial de spécification du modèle, celui où la connaissance géographique est injectée dans les agents et comportements, est celui des villes : les villes sont des entités indivisibles, ce sont les agents du modèle (niveau micro). Le niveau système de villes constitue le niveau émergent (niveau macro). SIMPOP_{NANO} travaille quant à lui au niveau inférieur : le niveau spatial de spécification est celui des fonctions et des quartiers (niveau micro), dans lesquels doivent se localiser *les mêmes fonctions urbaines* acquises par la ville au niveau système. Le niveau ville constitue le niveau émergent (niveau macro). Il ne s'agit donc pas de la « même » ville que celle de SIMPOP2 : celle de SIMPOP_{NANO} est une entité qui n'existe que dynamiquement, à travers une superposition de mesures qui permettent au géographe de décider si la structure émergente de son modèle ressemble ou non à une ville. Pour SIMPOP3 nous avons donc été confrontés au besoin de réconcilier ces deux représentations de la ville, et de construire un pont entre la ville spécifiée au niveau micro de SIMPOP2 et celle, émergente, au niveau macro de SIMPOP_{NANO}. Ce besoin est illustré par la figure 4.

Pour réaliser cette articulation entre les deux modèles, nous avons choisi d'identifier une description qui leur soit commune, qui représente à la fois un effet de la dynamique du système des villes sur la ville, et qui puisse exprimer une dynamique interne à la ville susceptible d'avoir des effets en retour sur le système. La notion de *fonction urbaine* a donc été instanciée dans le programme en agent « inter-niveaux ». La fonction urbaine est à l'origine une notion définie par les géographes pour qualifier le rôle spécifique joué par une ville au sein du système, et qui trouve donc sa pertinence au niveau du système de villes. Pour réaliser le couplage, la notion a été infléchie et l'hypothèse a été faite qu'elle pouvait être utilisée pour qualifier les grands types d'activités en interaction à l'intérieur de la ville au cours de son évolution. Ce choix se justifie par la période historique abordée par le modèle (1800-2000). C'est en fait cette extension d'un concept, du niveau inter-urbain au niveau intra-urbain, qui

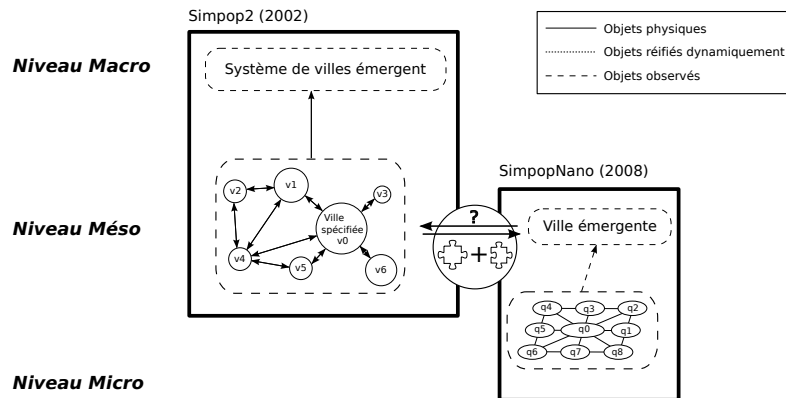


Figure 4. SIMPOP3 : Comment articuler l'agent ville spécifié de SIMPOP2 avec la ville émergente de SIMPOP3 ?

permet d'utiliser la *fonction urbaine* comme un agent « tampon » qui permet de faire le lien entre les règles des deux modèles, et de réaliser ce couplage.

4. Vers des organisations multi-agents multi-niveaux

Après avoir rendu compte de différentes expériences de modélisation à base d'agents à différents niveaux d'organisation, il s'agit maintenant d'essayer de les systématiser. L'objectif est de proposer une grille d'analyse de ces modèles multi-niveaux en vue de faciliter leur conception et leur implémentation.

4.1. Grille d'analyse

A partir des exemples présentés, ainsi qu'un ensemble d'autres modèles de la littérature, on peut tenter une catégorisation des modèles multi-niveaux en trois classes :

1) *Multi-modèles* : il s'agit du couplage de plusieurs modèles à l'origine indépendants. On considère des modèles à des niveaux différents en interaction. Les sorties des modèles de niveau supérieur servent à définir des paramètres globaux des modèles de niveau inférieur et les sorties des modèles de niveau inférieur sont considérées comme des flux d'entrée dans les modèles de niveau supérieur. On peut distinguer encore deux sous-catégories :

a) les différents modèles s'exécutent de manière alternée (c'est le cas par exemple de SIMPOP3). Dans ce cas, le couplage entre les modèles est généralement faible mais des difficultés peuvent néanmoins apparaître si les différents modèles partagent des objets qu'ils peuvent chacun modifier (l'entité ville dans le cas de SIMPOP3).

b) un modèle multi-agent est encapsulé à l'intérieur d'un agent de niveau supérieur. Si l'on pense au modèle de tumeurs, cela correspondrait au fait que l'on modélise la dynamique interne des cellules, plutôt que de considérer une production à taux fixe de molécules de PAI-1. Cela permettrait alors de prendre en compte les réseaux de régulation génique contrôlant la production de PAI-1. Il faut faire attention, dans ce cas, à l'ordonnancement de l'activité des agents aux différents niveaux, qui n'intervient pas nécessairement à la même échelle temporelle.

2) *Modèles à spécification multi-niveaux* : il s'agit de la cohabitation et de l'interaction, au sein d'un même modèle, d'agents situés à des niveaux différents le long des échelles spatiale, temporelle ou quantitative. On peut y parvenir d'au moins deux façons différentes :

a) par spécification statique des entités simulées (ex. : des cellules qui interagissent avec des molécules dans le modèle de tumeur ; des personnes qui interagissent avec des quartiers dans le modèle Accessim (Delâge *et al.*, 2008), SIMPOP3). A nouveau apparaissent des problèmes liés à l'ordonnancement des agents. Par exemple, les mécanismes de réaction de cellules et de molécules interviennent à des résolutions temporelles très différentes. Cela impose de privilégier un mode de simulation à évènements discrets, ou encore de gérer en parallèle plusieurs ordonnanceurs pour les différents types d'objets simulés.

b) par spécification dynamique des entités simulées : implémentation de mécanismes de type « loupe dynamique » qui permettent d'augmenter le degré de résolution spatiale sur des zones critiques. (ex. : le modèle fractal de l'environnement proposé par (Marilleau *et al.*, 2008)). La difficulté principale dans ce cas réside dans l'adaptation continue nécessaire pour toujours conserver le plus haut degré de détail sur les zones d'intérêt, par exemple celles où se trouvent des agents.

3) *Modèles multi-niveaux à réification dynamique* : il s'agit de modèles qui produisent des agents de niveau supérieur par « observation » automatique de la simulation, caractérisation de structures de plus haut niveau et réification de ces structures sous forme d'objets. On peut là encore distinguer deux cas distincts :

a) les agents des deux niveaux n'interagissent pas directement les uns avec les autres (ex. : simulBogota) car les entités des deux niveaux ne sont pas dans le même espace de modélisation. D'une certaine manière, ce cas ressemble au cas (1a), puisque l'on aura également une alternance entre l'exécution de modèles aux niveaux micro et macro : le simulateur exécute en effet en boucle l'évolution du modèle au niveau micro puis son analyse au niveau macro, puis l'évolution du modèle au niveau macro.

b) les agents des deux niveaux peuvent interagir entre eux et avec les agents de l'autre niveau : les entités des deux niveaux sont dans le même espace de modélisation (ex. : modèle de tumeur avec modèle agrégé, RIVAGE (Servat *et al.*, 1998), ou encore (Gaud *et al.*, 2008) qui se base sur des approches hiérarchiques de type holoniques). Cette approche, sans doute la plus riche, entraîne cependant une plus grande complexité puisqu'elle nécessite de pouvoir à la fois détecter l'apparition d'une structure émergente, caractériser sa dynamique et ses interactions avec les autres agents, vérifier que la simulation reste valide, et enfin vérifier les conditions de maintien de

la structure émergente. Il s'agit ainsi de contrôler de manière dynamique aussi bien la réification de structures émergentes à partir d'agents individuels, que leur disparition lorsque les conditions de maintien ne sont plus vérifiées.

4.2. Exemple

Dans cette partie nous présentons un modèle théorique des dynamiques de diffusion des fourmis de l'espèce *Linepithema humile*, communément connues sous le nom de *fourmis argentines* (Krusheľnycky, 2004). Notre objectif principal, qui est purement théorique, est de profiter de ce problème complexe et couvrant des niveaux d'échelles très étendus, aussi bien du point de vue spatial que temporel, pour montrer les possibilités d'utilisation unifiée des différentes approches de modélisation multi-niveaux présentées dans les sections précédentes.

Contexte et objectif

Les fourmis argentines appartiennent à une espèce très invasive qui a été véhiculée accidentellement par l'homme et qui a envahi et continue d'envahir des zones géographiques très vastes dans le monde entier. L'activité humaine a non seulement permis la diffusion de ces fourmis de continent en continent, mais elle joue également le rôle de principal vecteur de diffusion à des distances moyennes (Holway *et al.*, 2002). La force des *fourmis argentines* réside dans une capacité accrue de coopération entre ouvrières appartenant à des colonies différentes (*unicolonialisme*). Elle s'appuie également sur une diffusion de nouvelles colonies basée sur une stratégie de reproduction à l'abri des prédateurs et de progression uniforme sur des distances courtes. Lorsqu'une colonie atteint une taille importante, elle se divise en plusieurs colonies formées chacune par une reine et une partie des ouvrières de la colonie originelle. Les nouvelles colonies se placent à des distances relativement courtes par rapport à la colonie originelle (une unicolonie s'étend d'une distance comprise environ entre 15 m et 200 m par an). Cette stratégie de diffusion présente moins de risques que d'autres stratégies comme par exemple celle basée sur les vols nuptiaux pratiquée par les fourmis européennes.

Les fourmis collaborent en fonction de leur capacité de reconnaissance de leurs patrimoines génétiques respectifs. Si deux fourmis provenant de colonies différentes se reconnaissent comme génétiquement similaires elles peuvent collaborer, alors qu'elles s'agressent si elles se reconnaissent comme génétiquement différentes. Dans leur milieu naturel originel, la capacité qu'ont les fourmis de reconnaître le patrimoine génétique de leurs congénères est très développée, ce qui augmente les possibilités de confrontation entre fourmis provenant de colonies différentes et qui permet l'auto-régulation de la croissance de la population. Par contre dans les milieux que ces fourmis ont colonisés à cause de l'activité humaine, le mécanisme d'auto-régulation est beaucoup moins efficace principalement pour deux raisons :

1) lors des transports accidentels, très peu de fourmis ont été initialement déplacées ce qui induit une réduction de la diversité du patrimoine génétique ;

2) dans leurs nouveaux milieux et en l'absence de prédateurs, les tailles de colonies atteignent des valeurs très importantes. Ces tailles importantes sont corrélées à une diminution de la capacité individuelle de reconnaissance génétique (Giraud *et al.*, 2002).

Dans ces conditions, les unicolonies de fourmis argentines se sont répandues sur des zones géographiques très étendues formant des super-colonies (des ensembles d'unicolonies « alliées » s'étendant sur plusieurs milliers de kilomètres, comme par exemple la super-colonie méditerranéenne qui s'étend en continu le long de la côte depuis le sud de l'Italie jusqu'au Portugal (Giraud *et al.*, 2002)), entraînant au passage la disparition des espèces locales de fourmis et un déséquilibre très important dans les écosystèmes locaux (Holway *et al.*, 2002). Afin de contraindre ces invasions, certaines mesures visant à rétablir le mécanisme d'auto-régulation originel ont été proposées, comme par exemple l'introduction de la diversité génétique par l'implantation contrôlée de nouveaux allèles (Tsutsui *et al.*, 2001). Le modèle multi-niveaux théorique que nous présentons pourrait être utilisé pour évaluer les effets de ce type de mesures de contrôle sur la progression des super-colonies.

Modèle

Afin de bien saisir les dynamiques de diffusion et de concurrence entre unicolonies de fourmis argentines, il est nécessaire de représenter les processus de reconnaissance génétique et d'agression ainsi que les processus de diffusion, que celle-ci soit naturelle ou liée à l'intervention humaine. Nous proposons de considérer ces processus à trois niveaux de modélisation différents : au *niveau de la fourmi*, au *niveau de la colonie* et au *niveau de l'unicolonie*. La relation entre les différents processus et niveaux est décrite par la suite :

1) *La capacité de reconnaissance génétique* doit être modélisée au niveau de la colonie car elle est dépendante des variables inhérentes à ce niveau (taille de la colonie et variabilité génétique).

2) *L'agression* peut être modélisée aux niveaux des fourmis et des colonies. Dans le premier cas il s'agirait de considérer des combats entre fourmis dont le déclenchement dépendrait de la capacité individuelle de reconnaissance et du code génétique de chacune des fourmis. Dans le deuxième cas, il s'agirait d'une implémentation plus statistique du mécanisme d'agression qui dépendrait de calculs comme la probabilité de rencontre (dépendante de la proximité spatiale entre les colonies et de leurs tailles respectives) et la probabilité d'agression (dépendante de la capacité de reconnaissance et le vecteur de descripteurs génétiques moyens).

3) *La diffusion naturelle* peut être modélisée au niveau des colonies et au niveau des unicolonies. Dans le premier cas, une colonie qui atteint une certaine taille est capable de donner naissance à une nouvelle colonie à quelques mètres de distance par le déplacement d'un *groupe migratoire* (un groupe de fourmis composé d'une reine et d'un nombre minimum d'ouvrières³). Dans le deuxième cas une unicolonie se dif-

3. selon les observations, il suffirait d'une dizaine d'ouvrières

fuse de manière uniforme (de l'ordre de plusieurs dizaines de mètres par an). Dans l'absence de concurrence, cette diffusion ne peut être influencée que par la disponibilité de ressources dans l'environnement, la topographie, le climat et d'autres facteurs environnementaux.

4) *Dans la diffusion artificielle* le vecteur de diffusion est l'être humain⁴. Cette diffusion peut également être modélisée aux niveaux de la colonie et de l'unicolonie. Ce mécanisme pourrait être représenté sous la forme de flux de groupes migratoires sur le réseau des communications humaines.

Comme il a été noté, les différents mécanismes peuvent être implémentés à des niveaux différents de modélisation. Le choix du niveau dépend de la question à résoudre, de la granularité souhaitée et des ressources de calcul disponibles. Il est donc souhaitable de compter avec des mécanismes, automatiques ou manuels, de changement dynamique de granularité (si l'expérimentateur souhaite par exemple explorer une zone particulière). Dans le cas du mécanisme d'agression, si l'on s'intéresse à l'évolution de colonies voisines rivales, il est pertinent de passer au niveau des fourmis, mais si l'on s'intéresse à des dynamiques plus globales comme l'évolution de la frontière entre unicolonies rivales, considérer l'agression au niveau des fourmis serait trop coûteux du point de vue de la puissance de calcul, et il serait alors préférable de se placer au niveau des colonies.

Agents et niveaux

Pour modéliser intégralement l'évolution des super-colonies, il est nécessaire de coupler les différents niveaux proposés. Le *niveau des super-colonies* devrait émerger à partir des interactions entre les agents des autres niveaux. La super-colonie serait alors observée et prendrait la forme d'un réseau d'unicolonies alliées.

1) *Couplage entre les fourmis et les colonies* : comme présenté précédemment, le niveau des fourmis ne devrait être pris en compte que lorsque l'on s'intéresse à des phénomènes très localisés (concurrence entre deux colonies, augmentation manuelle du niveau de détail sur une zone particulière) en raison du coût important de calcul que cela implique. Dans ces cas, il n'est pas nécessaire d'avoir une représentation interne de la colonie car on ne s'intéresse qu'à l'action extérieure des fourmis (récolte de nourriture, combats). La colonie peut être vue comme un générateur de fourmis et les fourmis comme les fournisseurs de nutriments de la colonie. La capacité de création de fourmis dépendrait de la quantité de ressources reçue par la colonie.

2) *Couplage entre les colonies et les unicolonies* : pour effectuer ce couplage nous proposons d'implémenter des mécanismes de décomposition et de recomposition dynamiques des unicolonies en se basant sur leur distribution spatiale. D'un côté, *la décomposition d'une partie de l'unicolonie en colonies composantes* sera nécessaire dans les zones de contact entre unicolonies concurrentes. L'agression aura alors lieu

4. Les fourmis peuvent également être déplacées par des animaux de taille moyenne, par des ruisseaux, ou d'autres vecteurs de diffusion encore. Un raisonnement similaire peut alors être appliqué pour considérer ces différents vecteurs

entre colonies rivales sur les zones frontalières. Les colonies éloignées de ces zones frontalières peuvent continuer à être représentées par un seul *agent unicolonie*. D'un autre côté, *les colonies considérées de manière individuelle* et qui, en raison de la dynamique du système, se retrouvent éloignées des zones d'agression, *devront être réabsorbées par leur unicolonie (recomposition dynamique)*. Il faut également noter que l'on pourrait fusionner des unicolonies « alliées » qui entreraient en contact. Ces mécanismes de décomposition et de recomposition dynamique, qui permettraient un dosage adaptatif de la puissance de calcul, sont similaires à ceux proposés dans la modélisation de la croissance tumorale.

3) *Couplage entre l'unicolonie et la super-colonie* : on peut considérer un couplage basé sur l'échange de groupes migratoires entre les unicolonies grâce au réseau de diffusion. Les groupes migratoires qui se déplaceraient sur le réseau de diffusion seraient absorbés par (ou entreraient en conflit avec) les unicolonies locales dans les zones occupées, ou bien formeraient de nouvelles unicolonies dans les zones inoccupées. Cette forme d'articulation est similaire à celle proposée dans SIMPOP3 pour représenter les inter-influences entre le système de villes et la ville, via les fonctions urbaines qui, en étant modifiées par les dynamiques des deux niveaux, jouent le rôle d'un *tampon* permettant la communication entre modèles.

4) Enfin le *niveau des super-colonies* peut être construit dynamiquement afin de réaliser des simulations à large échelle et d'observer des interactions éventuelles entre unicolonies géographiquement éloignées. Pour construire ce niveau nous proposons d'utiliser une méthode basée sur la proximité génétique et les capacités de reconnaissance dans chaque unicolonie. Dans cette méthode, il s'agirait d'effectuer une classification automatique (similaire à celle qui est effectuée dans SIMULBOGOTA), basée sur une distance entre vecteurs génétiques de chacune des unicolonies et sur le calcul d'un seuil sur ces distances qui dépendrait des capacités de reconnaissance individuelles. Ces deux mesures gouverneraient alors la formation de « clusters d'unicolonies », qui correspondraient chacun à une super-colonie. Le comportement des unicolonies serait alors modifié du fait de leur appartenance à une super-colonie.

5. Conclusion

Nous avons présenté dans cet article plusieurs modèles à base d'agents qui peuvent tous, d'une manière ou d'une autre, être qualifiés de « multi-niveaux ». Ces modèles sont pourtant très différents les uns des autres, tant du point de vue de la taille des entités modélisées que des approches mises en oeuvre pour pallier les difficultés engendrées par le couplage de différents niveaux d'organisation. Après avoir clarifié les notions d'échelle et de niveau, nous avons souligné qu'il nous semblait plus juste de parler de « modèles multi-niveaux » que de « modèles multi-échelles ». Nous avons ensuite proposé une grille d'analyse permettant la classification de différents types de modèles multi-niveaux, à l'intérieur de laquelle nous avons positionné les trois modèles étudiés. Celle-ci demande à être complétée et raffinée en y intégrant d'autres travaux de la communauté.

Ce faisant, nous ne prétendons pas proposer une méthodologie de conception de modèles multi-niveaux, qui puisse être réutilisée, étape par étape, pour la conception d'un nouveau modèle. L'objectif est d'une part de montrer que différentes problématiques de modélisation peuvent conduire à des solutions très différentes et d'autre part d'aider le concepteur à identifier les points durs de chacune des approches. L'identification de ces difficultés nous a conduit à proposer quelques pistes de réalisation de couplage inter-niveaux et orientés agent. Leur pertinence a été établie pour les exemples étudiés, ainsi que pour d'autres issus de la littérature.

Lorsque des modèles existent dans un domaine particulier et qu'ils ont prouvé leur efficacité, il est à la fois utile et important de pouvoir les intégrer comme des composants dans des modèles de plus grande envergure. Cela implique de permettre le couplage de modèles à différents niveaux, ou fondés sur des formalismes hétérogènes (discret/continu par exemple). Pour ce faire, d'une part des modèles microscopiques peuvent être encapsulés sous formes d'agents, d'autre part des modèles continus (par exemple des modèles à base d'équations différentielles) peuvent servir dans la définition de variables et dynamiques globales utilisables comme entrées de modèles de niveau supérieur.

Parce que les systèmes complexes évoluent de manière dynamique, il est également important que le découpage en agents puisse, lui aussi, évoluer dynamiquement. Pour ce faire, il est nécessaire de permettre la détection (Moncion, 2008) et la réification automatique de structures émergentes, et réciproquement de permettre le découpage d'agents en organisations sous-jacentes, de manière interactive. En adaptant les niveaux de description du système, il ne s'agit pas tant de gagner en performances computationnelles, que de gagner en pouvoir explicatif en sélectionnant les niveaux d'organisation et de description pertinents, pour obtenir la plus grande compréhension possible des mécanismes à l'oeuvre dans le système complexe étudié.

6. Bibliographie

- Bertaud A., « The spatial organization of cities : deliberate outcome or unforeseen consequence ? », May, 2004, unpublished, available at http://alain-bertaud.com/images/AB_The_spatial_organization_of_cities_Version_3.pdf.
- Bura S., Guerin-Pace F., Mathian H., Pumain D., Sanders L., « Multi-agents system and the dynamics of a settlement system », *Geographical Analysis*, vol. 28, n° 2, p. 161-178, 1996.
- Delâge M., Le Néchet F., Louail T., « ACCESSIM : modélisation et simulation d'accessibilité et d'inégalités géographiques dans la ville », *Mappemonde*, 2008. <http://mappemonde.mgm.fr/num20/fig08/fig08405.html>.
- Gaud N., Galland S., Gechter F., Hilaire V., Koukam A., « Holonic multilevel simulation of complex systems : Application to real-time pedestrians simulation in virtual urban environment », *Simulation Modelling Practice and Theory*, vol. 16, n° 10, p. 1659-1676, 2008.
- Gibson C. C., Ostrom E., Ahn T. K., « The concept of scale and the human dimensions of global change : a survey », *Ecological Economics*, vol. 32, n° 2, p. 217-239, 2000.

- Gignoux J., Davies I., Hill D., « 3Worlds : a new platform for simulating ecological systems. », *1st Open Inter. Conf. on Modelling and Simulation*, Clermont-Ferrand, p. 49-64, 2005.
- Gil-Quijano J., Piron M., « Formation automatique de groupes d'agents sociaux par techniques d'apprentissage non supervisé », *Actes de EGC'07 Atelier Fouille de Données et Algorithmes Biomimétiques*, Namur Belgique, 23 Janvier, 2007a.
- Gil-Quijano J., Piron M., Drogoul A., *Mechanisms of automated formation and evolution of social-groups : A multi-agent system to model the intra-urban mobilities of Bogotá city*, Idea Group Inc., chapter 12, p. 151-168, 2007b.
- Giraud T., Pedersen J. S., Keller K. L., « Evolution of supercolonies : The Argentine ants of southern Europe », *PNAS*, vol. 99, p. 6075-6079, 2002.
- Holway D. A., Lach L., Suarez A. V., Tsutsui N. D., Case T. J., « The causes and consequences of ant invasions », *Annu. Rev. Ecol. Syst.*, vol. 33, p. 181-233, 2002.
- Krushelnicky P., Harmful Non-Indigenous Species Report on Argentine ant, Technical report, Hawaiian Ecosystems at Risk project (HEAR), Haleakala Field Station of USGS (PIERC), 2004.
- Lepagnot J., Hutzler G., « A multi-scale agent-based model for the simulation of avascular tumor growth », *Jour. of Biol. Phys. and Chem.*, vol. 9, p. 17-25, March, 2009.
- Lesne A., « Biologie des systèmes : l'organisation multiéchelle des systèmes vivants », *Medecine Sciences*, vol. 25, p. 585-587, 2009.
- Louail T., « De la géométrie des réseaux viaires à l'organisation spatiale intra-urbaine : une approche de la comparaison des formes de villes par la simulation à base d'agents », *Revue Internationale de Géomatique*, vol. 19, n° 4, p. 443-469, 2009.
- Maquerlot F., Galiacy S., Malo M., Guignabert C., Lawrence D. A., d'Ortho M.-P., Barlovatz-Meimon G., « Dual Role for Plasminogen Activator Inhibitor Type 1 as Soluble and as Matricellular Regulator of Epithelial Alveolar Cell Wound Healing », *Am. J. Pathol.*, vol. 169, p. 1624-1632, 2006.
- Marilleau N., Cambier C., Drogoul A., Perrier E., Chotte J., Blanchart E., « Environnement multi-échelles à base de fractales pour la modélisation agent d'écosystèmes », *JFSMA'08*, Brest, 2008.
- Moncion T., Modélisation de la complexité et de la dynamique des simulations multi-agents. Application pour l'analyse des phénomènes émergents, Thèse de doctorat, Université d'Evry, 2008.
- Pumain D., « Pour une théorie évolutive des villes », *L'Espace Géographique*, vol. 2, p. 119-134, 1997.
- Pumain D., Sanders L., Bretagnolle A., Glisse B., Mathian H., « The Future of Urban Systems », in D. Lane, D. Pumain, S. Van der Leeuw, G. West (eds), *Complexity perspectives on innovation and social change*, Methodos Series, Springer, chapter 12, 2009.
- Servat D., Perrier E., Treuil J.-P., Drogoul A., « When Agents Emerge from Agents : Introducing Multi-scale Viewpoints in Multi-agent Simulations », in J. Sichman, R. Conte, N. Gilbert (eds), *MABS*, vol. 1534 of *LNCs*, Springer, p. 183-198, 1998.
- Treuil J.-P., Drogoul A., Zucker J.-D., *Modélisation et simulation à base d'agents*, Dunod, 2008.
- Tsutsui N. D., Case T. J., « Population genetics and colony structure of the Argentine ant (*Linepithema humile*) in its native and introduced ranges », *Evolution*, vol. 55, p. 976-985, 2001.